最小カット問題に対する Kargerのランダムアルゴリズムの新しい確率的評価について

戴 陽* 岩野 和生† 加藤 直樹*

* 神戸商科大学 管理科学科
† 日本IBM東京基礎研究所

G を n 個の節点をもつ無向グラフとする。Karger はグラフの最小カット問題に対してランダムアルゴリズムを与えた。そのアルゴリズムは確率$\Omega(n^{-2})$ でグラフのある特定の最小カットを見つける。本論文は Karger のアルゴリズムに対して新しい確率的評価を与える。ランダムグラフの特定の最小カットを見つける確率の期待値は$\Omega(p/n)$ であることを示す。ただし、p は枝を選ぶ確率である。さらに、密な一般グラフに対しても新しい結果を示す。

A New Probabilistic Evaluation of Karger's Randomized Algorithm for Minimum Cut Problems

Yang Dai* Kazuo Iwano† Naoki Katoh*

* Department of Management Science, Kobe University of Commerce, Gakuen-Nishimachi 8-2-1, Nishi-ku, Kobe 651-21, Japan, dai@kucgw.kobeuc.ac.jp., naoki@kucgw.kobeuc.ac.jp.

† Tokyo Research Laboratory, IBM Japan, 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 242, Japan, iwano@trl.vnet.imb.com.

ABSTRACT: Recently Karger proposed a new randomized algorithm for finding a minimum cut of an n-vertex graph with probability $\Omega(n^{-2})$. In this paper we present a new probabilistic evaluation of Karger's randomized algorithm. For random graphs whose edges are selected with a given probability p, $\frac{\log n}{n} \leq p \leq 1$, we show that the expectation of success probability of the algorithm is $\Omega(\frac{p}{n})$. We also investigate a class of graphs with special structure that consists of two n-cliques and $\gamma(n-1)$ edges between the two cliques. Here γ is a parameter satisfying $0 < \gamma < 1$ that makes $\gamma(n-1)$ edges a unique minimum cut. We show that the algorithm finds the unique minimum cut with probability $\Omega(\frac{\gamma}{n^\gamma})$. A tighter bound is also given for general dense graphs.

1 Introduction

We give a new probabilistic evaluation of Karger's newly proposed randomized algorithm [6] for minimum cut problems. Given a connected undirected simple graph $G = (V, E)$ with n vertices and m edges, the minimum cut problem is to find a set of edges of minimum cardinality whose removal leaves two connected components. The minimum cut problem is extensively studied. For undirected graphs, the current deterministic fastest algorithm requires $O(m\lambda(G)\log(n^2/m))$ time [4,5], where $\lambda(G)$ is the edge connectivity of G.

Recently, Karger [6] proposed a new randomized algorithm (called contraction algorithm below) for computing minimum cuts based on a simple operation of edge contraction. When the contraction algorithm terminates, a minimum cut is produced with probability $\Omega(n^{-2})$. So, if we perform $O(n^2 \log n)$ independent trials, we find a minimum cut with high probability. On the other hand, Dai et al. [2] devised a heuristic approach to approximately solve a minimum cut problem by using an efficient algorithm for minimum range cut problems [8]. The approach is based on an association of range value of a cut and its cut value when each edge weight is chosen uniformly randomly from a fixed interval.

Although these two approaches are apparently different, they are closely related to each other. It is known that at each run of minimum range cut algorithm the heuristic method finds a cut corresponding to the one obtained by a single trial of Karger's algorithm [2]. It was shown that the method finds a minimum cut with a probability higher than or equal to the one that contraction algorithm does at a single trial [2]. In [2] the usefulness of the heuristic method is demonstrated by computational results, which report that solutions obtained within only $\sqrt{n}$ iterations are very close to the optimal for several types of graphs, such as graphs generated by NETGEN [9] and other random graphs [3]. Moreover, from our computational results the contraction algorithm produces minimum cuts with probability much higher than $\Omega(n^{-2})$ for most of the input instances. On the other hand, Karger's bound is tight for unicycle graphs. This implies a gap between theoretical and computational results for random graphs. We shall give a bound stronger than $\Omega(n^{-2})$ on the probability that these two methods succeed for random inputs. A stronger bound will give a theoretical support on the effectiveness of both the heuristic and contraction algorithms.

The key point of analysis of the contraction algorithm is to give an upper bound of the number of edges contracted during the process of edge contraction. By introducing a scheme called contraction tree we obtain new upper bounds of the number for the following several types of graphs.

Complete graphs: The contraction algorithm finds a specific minimum cut with probability $\Omega(n^{-1})$ for complete graphs.

Graphs with two cliques: Each graph contains two n-cliques and $\gamma(n-1)$ edges between the two cliques, where γ is a parameter satisfying $0 < \gamma < 1$ that makes the graph have a unique minimum cut separating the two cliques. We show that the contraction algorithm finds the unique minimum with probability $\Omega(\frac{\gamma}{n^\gamma})$.

Random graphs: The random graphs are generated by a model introduced in [10] i.e., given a probability $0 < p < 1$, each random graph is generated by choosing edges between vertices with probability p. Therefore each random graph G is generated with certain probability q_G. For given p and n, we define $G(n,p)$ as all simple graphs with n vertices with which q_G is associated. We assume $\frac{\log n}{n} \leq p \leq 1$ guaranteeing that almost all random graphs generated are connected [10]. When performing the contraction algorithm on a given instance G of the random graphs, let $P_G(\text{surv})$ denote the probability that Karger's contraction algorithm obtains a specific minimum cut for G. We shall show that the expectation of the probability $P_G(\text{surv})$ for random graphs G is $\Omega(\frac{p}{n})$.

General dense graphs: We consider graphs with n vertices and m edges. Define the density of a graph as $\alpha = \frac{2m}{n(n-1)}$. For a graph with density $(1 - \frac{1}{\sqrt{n}})^2 < \alpha \leq 1$, we show that the contraction algorithm finds a specific minimum cut with probability $\Omega(\frac{1}{\sqrt{\alpha} n^{1+\rho}})$, where $\rho = \dfrac{\log(1 + n^2(1-\sqrt{\alpha})^2)}{\log n}$. Since $(1 - \frac{1}{\sqrt{n}})^2 < \alpha \leq 1,\ 1 > \rho \geq 0$. Therefore the above bound is tighter than the bound given by Karger [6]. (The derivation of this result is omitted in this version.)

2 A new probabilistic evaluation for random graphs

The contraction algorithm proposed by Karger [6] is very simple. The algorithm repeats one fundamental operation: contracting a randomly chosen edge in a graph. Contraction of an edge connecting two vertices v_1 and v_2 results in a new vertex v called *super vertex*, by removing all edges between v_1 and v_2 and letting the set of edges incident on v be the union of the sets of remaining edges incident on v_1 and v_2. Karger gives a probabilistic evaluation based on the following observation. Take some specific minimum cut of c edges. Suppose that we never select any edge in the minimum cut during contractions. Then after each contraction the contracted graph must have a minimum degree of at least c, since the contracted graph has a minimum cut of c edges. Let $m(j)$ denote the number of remaining edges at the beginning of the jth contraction. Since the number of vertices decreases by one for each contraction, we have

$$m(j) \geq (n - j + 1)c/2. \tag{1}$$

2.1 A worst-case analysis for a complete graph and its variant

We introduce a new scheme to obtain a different way of counting the number of edges in contracted graphs, which provides a more accurate estimation. We represent a sequence of contractions by a binary tree called a *contraction tree*. Each leaf of the contraction tree corresponds to a vertex in the original graph. When edges between two super vertices in the contracted graph are contracted, we create an internal node which has two children in the contraction tree corresponding to the two super vertices.

In order to clearly illustrate our new technique, we shall first analyze the performance of the contraction algorithm when applied to complete graphs. algorithm on the graph. Notice that a minimum cut separates one node from the other nodes, and there are n distinct minimum cuts in the complete graph with n vertices. Consider a particular minimum cut, and suppose that the contraction algorithm never contracts any edge in the minimum cut during the course of the alogrithm. Let $d(j)$ denote $d(j)$ is the sum of number of edges contracted during the first j contractions. We have the following lemma.

Lemma 1 $d(j) \leq j(j+1)/2$.

Proof. We prove it by induction on j. For $j = 1$ the result holds since the number of contracted edges is one. Assume that the result holds for $j-1$. Now we consider the case of j. Suppose that the jth contraction creates an internal node whose left subtree is T_l and right subtree is T_r. Suppose that T_l (*resp.* T_r) has q_1 (*resp.* q_2) leaves. From the induction hypothesis, the number of contractions performed corresponding to T_l (*resp.* T_r) is at most $q_1(q_1-1)/2$ (*resp.* $q_2(q_2-1)/2$) edges. Notice that in order to contract q_1 (*resp.* q_2) vertices, we need $q_1 - 1$ (*resp.* $q_2 - 1$) contractions. Moreover, the number of edges contracted by the jth contraction is $q_1 q_2$. In general, there may be other internal nodes which are not in the subtrees of T_r and T_l. For simplicity we assume that all these internal nodes form a single subtree containing q_3 leaves. The other cases can be analyzed by a similar argument as follows. We have

$$d(j) \leq q_1(q_1-1)/2 + q_2(q_2-1)/2 + q_1 q_2 + q_3(q_3-1)/2.$$

Since the current contraction is the jth contraction,

$$(q_1-1)+(q_2-1)+(q_3-1)+1 = q_1+q_2+q_3-2 = j.$$

Thus we have

$$\begin{aligned} d(j) &\leq q_1(q_1+q_2-1)/2 + q_2(q_1+q_2-1)/2 + q_3(q_3-1)/2 \\ &= (q_1+q_2)(q_1+q_2-1)/2 + q_3(q_3-1)/2 \\ &\leq (q_1+q_2)(q_1+q_2-1)/2 + q_3(q_3-1)/2 + (q_1+q_2-1)(2q_3-2)/2 \\ &= (q_1+q_2+q_3-2)(q_1+q_2+q_3-1)/2 = j(j+1)/2. \end{aligned}$$ ■

By Lemma 1 the maximum value of $d(j)$ is realized when each contraction is carried out in such a way that a single vertex and a super vertex are contracted into a new super vertex, and the number of edges contracted is at most $j(j+1)/2$ during the first j contractions. Therefore

$$m(j) = m - d(j-1) \geq \frac{1}{2}(n(n-1) - j(j-1)). \tag{2}$$

Notice that for an n-vertex complete graph the value of $m(j)$ in (2) is greater than the value of $m(j)$ in (1). Let $P_j(\text{dest})$ denote the conditional probability that some specific minimum cut is destroyed for the first time at the jth contraction on the condition that it has not been destroyed for the first $j-1$ contractions. Since a minimum cut has $n-1$ edges in a n-vertex complete graph,

$$P_j(\text{dest}) = \frac{n-1}{m(j)} \leq \frac{2(n-1)}{n(n-1) - j(j-1)}.$$

Let $P_j(\text{surv})$ be the conditional probability that some specific minimum cut survives at jth contraction, on the condition that it survives after the first $j-1$ contractions. We have

$$P_j(\text{surv}) = 1 - P_j(\text{dest}) \geq 1 - \frac{2(n-1)}{n(n-1) - j(j-1)}.$$

Let $P(\text{surv})$ denote the survival probability of some specific minimum cut after $(n-2)$ contractions.

Theorem 2 *The contraction algorithm produces a specific minimum cut of an n-vertex complete graph with probability* $\Omega(n^{-1})$.

Proof. When the contraction algorithm terminates, each original vertex has been merged into one of the two remaining super vertices. This obviously defines a cut of the original graph. Thus,

$$P(\text{surv}) = \prod_{j=1}^{n-2} P_j(\text{surv}) \geq \prod_{j=1}^{n-2} \left(1 - \frac{2(n-1)}{n(n-1) - j(j-1)}\right) = \prod_{j=1}^{n-2} \frac{(n-j-1)(n+j-2)}{(n-j)(n+j-1)} = \frac{1}{2n-3}. \quad \blacksquare$$

Corollary 3 *The contraction algorithm produces a minimum cut of an n-vertex complete graph with probability higher than* $1/2$.

The discussion made above can be extended to a variant of complete graphs. Let G be a graph of $2n$ vertices that has a special structure of containing two n-cliques and a unique minimum cut separating them. Assume that the number of edges in the minimum cut is $\gamma(n-1)$, where $0 < \gamma < 1$. The graph has $m = n(n-1) + \gamma(n-1)$ edges. Suppose that J contractions have been executed on the graph and no edges of the minimum cut are contracted. We shall obtain the upper bound of the number of edges contracted when the contraction algorithm for this graph is carried out until J contractions are performed. It is shown in the same manner as in Lemma 1 that (i) when $J < n$, the most destructive way of contraction is to consecutively contract the edges only within one of the cliques in such a way that every cotraction contracts an edge connecting a single vetex and a super vertex, and that (ii) when $n \le J \le 2n-2$, the most destructive way of contraction is to first consecutively contract the edges in one clique in a manner as described in (i) until the clique is contracted into a single super vertex, and then contract edges in the other clique again in the manner as described in (i). Thus the number of edges remaining at the beginning of the jth contraction satisfies

$$m(j) \ge \begin{cases} n(n-1) - \frac{1}{2}j(j-1) + \gamma n & j \le n-1 \\ \frac{1}{2}n(n-1) - \frac{1}{2}(j-n)(j-n+1) + \gamma n & \text{otherwise.} \end{cases}$$

The following lemma is required to prove our desired results.

Lemma 4 *Let A, B and C be given positive numbers with $A \ge B$. Then*

$$\frac{B+C}{A+C} \ge \frac{B}{A}. \tag{3}$$

The survival probability of the minimum cut is

$$\begin{aligned} P(\text{surv}) &\ge \prod_{j=1}^{n-1} \frac{n(n-1) - \frac{1}{2}j(j-1)}{n(n-1) - \frac{1}{2}j(j-1) + \gamma(n-1)} * \prod_{j=n}^{2n-2} \frac{\frac{1}{2}n(n-1) - \frac{1}{2}(j-n)(j-n+1)}{\frac{1}{2}n(n-1) - \frac{1}{2}(j-n)(j-n+1) + \gamma(n-1)} \\ &= \prod_{j=1}^{n-1} \frac{2n(n-1) - j(j-1)}{2n(n-1) - j(j-1) + 2\gamma(n-1)} * \prod_{i=1}^{n-1} \frac{n(n-1) - i(i-1)}{n(n-1) - i(i-1) + 2\gamma(n-1)}. \end{aligned} \tag{4}$$

We derive lower bounds of the two parts in (4) as follows.

$$\begin{aligned} & \prod_{j=1}^{n-1} \frac{2n(n-1) - j(j-1)}{2n(n-1) - j(j-1) + 2\gamma(n-1)} \\ (\text{replace } \gamma \text{ by } n/(n-1)) \quad &> \prod_{j=1}^{n-1} \frac{(\sqrt{2}n + j - 1)(\sqrt{2}n - j - 1) + (j + 2\sqrt{2}n - 2n - 1)}{(\sqrt{2}n + j)(\sqrt{2}n - j) + j} \\ (\text{by (3)}) \quad &> \frac{\sqrt{2}n}{(1+\sqrt{2})^2(\sqrt{2}n - 1)}. \end{aligned}$$

We use Gamma function [1] to evaluate the second term of the last equation in (4).

$$\begin{aligned} \prod_{i=1}^{n-1} \frac{n(n-1) - i(i-1)}{n(n-1) - i(i-1) + 2\gamma(n-1)} &= \prod_{i=1}^{n-1} \frac{(n-i)(n+i-1)}{(n-i+\gamma)(n+i-1+\gamma) - (\gamma + \gamma^2)} \\ &> \prod_{i=1}^{n-1} \frac{(n-i)(n+i-1)}{(n-i+\gamma)(n+i-1+\gamma)} = \Theta(\frac{\gamma}{n^\gamma}). \end{aligned}$$

Theorem 5 *For graphs described above, the contraction algorithm produces the unique minimum cut with probability $\Omega(\frac{\gamma}{n^\gamma})$ for $0 < \gamma < 1$.*

2.2 Asymptotic analysis for random graphs

In this section we analyze the asymptotic performance of the algorithm for random graphs. The random graph model used here is the one described in [10], which was given in Section 1. The sample space, denoted by $S(G(n,p))$, consists of all labeled graphs G defined above for $\frac{\log n}{n} \leq p \leq 1$. Each graph in $S(G(n,p))$ is generated with certain probability, and a graph in $G(n,p)$ is expected to have $pn(n-1)/2$ edges. It is obvious that with high probability the number c of edges in a minimum cut of a graph satisfies

$$c \leq p(n-1). \tag{5}$$

Let $Ex(\cdot)$ denote the expectation of a random variable associated with a graph from $G(n,p)$. The following lemma is proved in a manner similar to Lemma 1.

Lemma 6 *Let $Ex(d(j))$ be the expectation of the number of edges contracted during the first j contractions. Then*

$$Ex(d(j)) \leq \frac{pj(j-1)}{2} + j. \tag{6}$$

Note that when $p = 1$ Lemma 6 coincides with that of Lemma 1. By the same argument in the previous section the expected number of remaining edges before the jth contraction is

$$Ex(m(j)) = Ex(m - d(j-1)) \geq \frac{pn(n-1)}{2} - \frac{p(j-2)(j-1)}{2} - (j-1). \tag{7}$$

When performing the contraction algorithm on a given instance G of the random graphs, we define the conditional probability $P_{Gj}(\text{surv})$ that a specific minimum cut of G survives at the jth contraction, on the condition that no edges in the cut has been contaracted until first j contractions. Let $Ex(P_{Gj}(\text{surv}))$ denote the expectation of $P_{Gj}(\text{surv})$ taken over sample space $S(G(n,p))$. Similarly let $Ex(P_G(\text{surv}))$ be the expectation of the probability that the algorithm succeeds in finding a specific minimum cut of G. Therefore the expectation $Ex(P_G(\text{surv}))$ is given by $\sum_{G \in G(n,p)} P_G(\text{surv}) q_G$.

$$\begin{aligned}
Ex(P_{Gj}(\text{surv})) &\geq 1 - \frac{2p(n-1)}{pn(n-1) - p(j-2)(j-1) - 2(j-1)} \\
&\geq \frac{(n-2)(n-1) - (j-2)(j-1) - 2(j-1)/p}{n(n-1) - (j-2)(j-1) - 2(j-1)/p} \\
&= \frac{(n+j-1+1/p)(n-j-2-1/p) + (4j+3/p+1/p^2-2)}{(n+j+1/p)(n-j-1-1/p) + (4j+3/p+1/p^2-2)} \\
\text{(by (3))} \quad &\geq \frac{(n+j-1+1/p)(n-j-2-1/p)}{(n+j+1/p)(n-j-1-1/p)}.
\end{aligned} \tag{8}$$

Since when $j \geq n - 3 - \lceil 2/p \rceil$ Karger's lower bound of (1) becomes better than (8), we have the following by letting $J = n - 3 - \lceil 2/p \rceil$:

$$Ex(P_G(\text{surv})) \geq \prod_{j=1}^{J} Ex(P_{Gj}(\text{surv})) * \prod_{j=J+1}^{n-2} \left(1 - \frac{2}{n-j+1}\right) \tag{9}$$

$$\geq \prod_{j=1}^{J} \frac{(n+j-1+1/p)(n-j-2-1/p)}{(n+j+1/p)(n-j-1-1/p)} * \prod_{j=J+1}^{n-2} (1 - \frac{2}{n-j+1})$$
$$= \frac{(n-J-2-1/p)(n+1/p)}{(n-2-1/p)(n+J+1/p)} * \frac{2}{(n-J)(n-J-1)} = \Theta(\frac{p}{n}).$$

Theorem 7 *For random graphs in $G(n,p)$, $\frac{\log n}{n} \leq p \leq 1$, the expectation of the probability that the contraction algorithm succeeds in finding a specific minimum cut is $\Omega(\frac{p}{n})$.*

In section 2.1 we considered a variation of the complete graph, i.e., a graph with two cliques of same size. We can apply the discussion to a more general graph, which is described in [3]. Let the graph have $2n$ vertices. The vertex set is randomly partitioned into two subsets with equal size. Edges joining vertices in the same subset are selected with probability p and edges joining vertices in different subsets are selected with probability $p\gamma/(n-1)$, $0 < \gamma < 1$. The cut separating the two subsets is almost surely a unique cut of the graph . Combine the discussion of this and the previous sections it is not difficult to obtain an expectation of probability $\Omega(\gamma p/n^{\gamma})$ that the contraction algorithm finds the desired cut.

References

[1] Dieudonne, J., *Infinitesimal Calculus*, Hermann, 1971.

[2] Dai, Y., H. Imai, K. Iwano, K. Katoh, K. Ohtsuka, and N. Yoshimura, A new unified approximate approach to the minimum cut problem and its variants using minimum range cut algorithms, IBM Research report, RT0082 (1992). (to appear in *Discrete Applied Math.*)

[3] Dyer, M.E. and A.M. Frieze, The solution of some random NP-hard problems in polynomial expected time, *Journal of Algorithms* 10 (1989), 452-489.

[4] Gabow, H.N., A matroid approach to finding edge connectivity and packing arborescences, *the Proc. ACM 23rd Symp. Theoretical Computer Science*, (1991), pp. 112-122.

[5] Gabow, H.N., Applications of a poset representation to edge connectivity and graph rigidity, *Proc. 32nd IEEE Symp. Found. Compt. Sci.*, (1991), pp. 812-821.

[6] Karger, D.R., Global min-cuts in RNC, and other ramifications of a simple min-cut algorithm, *Proc. 3rd ACM/SIAM Symp. Discrete Algorithms*, (1993), pp. 21-31.

[7] Karger, D.R., Random sampling in matriods, with applications to graph connectivity and minimum spanning trees, *Proc. 34th IEEE Symp. Found. Compt. Sci.*, (1994), pp. 84-93.

[8] Katoh, N. and K. Iwano, Efficient algorithms for minimum range cut problems, IBM Research Report, *Proceeding of 2nd WADS, Lecture Notes in Computer Science*, 519, Springer-Verlag, (1991), pp. 80-91. (to appear in *Networks*).

[9] Klingman, D., A. Napier, and J. Stutz, NETGEN: A program for generating large scale capacitated assignment, transportation, and minimum cost flow network problems, *Management Science* 20 (1974), 814-821.

[10] Palmer, E.M., *Graphical Evolution*, John Wiley & Sons, 1985.